(c) It can be shown that

$$\log_2\left(\sum_x p_r(x)\,c_r(x)\right) \le C \le \log_2\left(\max_x c_r(x)\right).$$

- If the lower-bound and upper-bound above are close enough. We take $p_r(x)$ as our answer and the corresponding capacity is simply the average of the two bounds.
- Otherwise, we compute the pmf

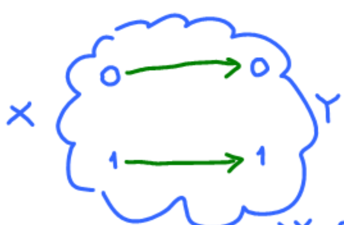
$$p_{r+1}(x) = \frac{p_r(x)\,c_r(x)}{\sum\limits_x p_r(x)\,c_r(x)} \quad \text{for all } x \in \mathcal{X}$$

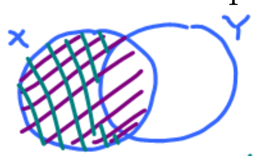
and repeat the steps above with index r replaced by $r+1$.

4.4 Special Cases for Calculation of Channel Capacity

In this section, we study special cases of DMC whose capacity values can be found (relatively) easily.

Example 4.27. Continue from Example 4.8 where we considered a noiseless binary channel. Find the corresponding channel capacity.





$$Q = \begin{array}{c|cc} x \backslash y & 0 & 1 \\ \hline 0 & 1 & 0 \\ 1 & 0 & 1 \end{array}$$

$I(X;Y) = H(X) - \cancel{H(X|Y)}^{\,0}$

$H(X|Y) = 0$

$\| $

$\sum_y q(y)\underbrace{H(X|Y=y)}_{=0}$

so, for this example, to maximize $I(X;Y)$, we need maximize $H(X)$ by using uniform $\underline{p} = [1/2 \;\; 1/2]$

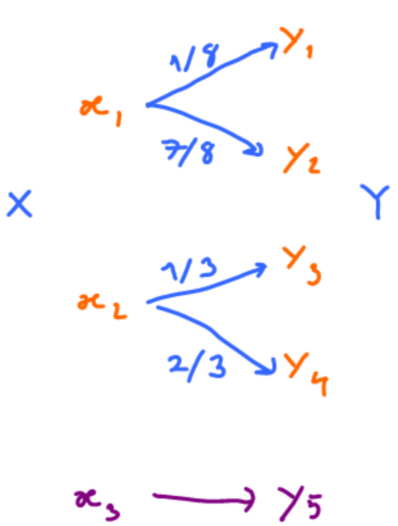
$C = \log_2|\mathcal{X}| = \log_2 2 = 1$ [bpcu]

bit per channel use

Example 4.28. Noisy Channel with Nonoverlapping Outputs: Find the channel capacity of a DMC whose

Example 4.28.2

$$\mathbf{Q} = \begin{array}{c|ccccc} x \backslash y & y_1 & y_2 & y_3 & y_4 & y_5 \\ \hline x_1 & 1/8 & 7/8 & 0 & 0 & 0 \\ x_2 & 0 & 0 & 1/3 & 2/3 & 0 \\ x_3 & 0 & 0 & 0 & 0 & 1 \end{array}$$



$I(X;Y) = H(X) - \underbrace{H(X|Y)}_{0}$

Again, $H(X|Y=y) = 0$ for all y.

So, $H(X|Y) = 0$.

To maximize $I(X;Y)$, we need to maximize $H(x)$ by uniform

$\underline{p} = [1/2 \;\; 1/2]$

$\underline{p} = [1/3 \;\; 1/3 \;\; 1/3]$

$C = \log_2|\mathcal{X}| = \log_2 2 = 1$ [bpcu]

$C = \log_2 3$ [bpcu]

In this example, the channel appears to be noisy, but really is not. Even though the output of the channel is a random consequence of the input, the input can be determined from the output, and hence every transmitted bit can be recovered without error.

4.29. Reminder:

(a) Some definitions involving entropy

(i) Binary entropy function: $h(p) = -p\log_2 p - (1-p)\log_2(1-p)$

(ii) $H(X) = -\sum_x p(x)\log_2 p(x)$

(iii) $H(\underline{\mathbf{p}}) = -\sum_i p_i \log_2(p_i)$

(b) A key entropy property that will be used frequently in this section is that for any random variable X,

$$H(X) \le \log_2 |\mathcal{X}| \text{ with equality iff } X \text{ is uniform.}$$

4.30. A DMC is a **noisy channel with nonoverlapping outputs** (NO^2) when there is only one non-zero element in each column of its $\mathbf{Q}$ matrix. For such channel,

$$C = \log_2 |\mathcal{X}| \text{ is achieved by uniform } p(x).$$

Definition 4.31. A DMC is called **symmetric** if (1) all the rows of its probability transition matrix $\mathbf{Q}$ are permutations of each other **and** (2) so are the columns.

Example 4.32. For each of the following $\mathbf{Q}$, is the corresponding DMC symmetric?

$$\begin{bmatrix} 0.3 & 0.2 & 0.5 \\ 0.5 & 0.3 & 0.2 \\ 0.2 & 0.5 & 0.3 \end{bmatrix}, \quad \begin{bmatrix} 0.3 & 0.2 & 0.5 \\ 0.5 & 0.2 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{bmatrix}, \quad \begin{bmatrix} 1/3 & 1/6 & 1/2 \\ 1/3 & 1/2 & 1/6 \end{bmatrix}, \quad \begin{bmatrix} 0.1 & 0.9 \\ 0.4 & 0.6 \end{bmatrix}$$

Symmetric?	① ✓ ② ✓ } ⇒ Yes	① ✓ ② × } ⇒ No	① ✓ ② × } ⇒ No	① × ② × } ⇒ No
Weakly symmetric?	(1) ✓ (2) ✓ } ⇒ Yes	(1) ✓ (2) × } ⇒ No	(1) ✓ (2) ✓ } ⇒ Yes	(1) × (2) × } ⇒ No

4.33. Q: Does symmetric DMC always have square $\mathbf{Q}$?

A: No

$$I(X;Y) = H(X) - \overbrace{H(X|Y)}^{\neq 0}$$

Example 4.34. Find the channel capacity of a DMC whose

$$\mathbf{Q} = \begin{array}{c} \\ x_1 \\ x_2 \\ x_3 \end{array}\begin{array}{c} \begin{array}{ccc} y_1 & y_2 & y_3 \end{array} \\ \begin{bmatrix} 0.3 & 0.2 & 0.5 \\ 0.5 & 0.3 & 0.2 \\ 0.2 & 0.5 & 0.3 \end{bmatrix} \end{array}.$$

$\underline{r}$

Solution: First, recall that the capacity C of a given DMC can be found by (34):

$$C = \max_{\underline{\mathbf{p}}} I(X;Y) = \max_{\underline{\mathbf{p}}} I(\underline{\mathbf{p}}, \mathbf{Q}).$$

$$H(\overbrace{[0.3\ 0.2\ 0.5]}^{\underline{r}}) \approx 1.4855$$

$$I(X;Y) = H(Y) - H(Y|X) = \sum_x p(x) \overbrace{H(Y|X=x)} = H(\underline{r})$$

We see that, to maximize $I(X;Y)$, we need to maximize $H(Y)$. Of course, we know that the maximum value of $H(Y)$ is $\log_2 |\mathcal{Y}|$ which happens when Y is uniform. Therefore, if we can find $\underline{\mathbf{p}}$ which makes Y uniform, then this same $\underline{\mathbf{p}}$ will give the channel capacity.

$$\underline{q} = \underline{p}\mathbf{Q}$$

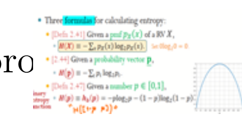
$$\left[\tfrac{1}{3}\ \tfrac{1}{3}\ \tfrac{1}{3}\right] = \left[\tfrac{1}{3}\ \tfrac{1}{3}\ \tfrac{1}{3}\right] \begin{bmatrix} 0.3 & 0.2 & 0.5 \\ 0.5 & 0.3 & 0.2 \\ 0.2 & 0.5 & 0.3 \end{bmatrix}$$

Try uniform X

Get uniform Y

$$C = \log_2 |\mathcal{Y}| - H(\underline{r}) \approx \log_2 3 - 1.4855 \approx 0.0995 \text{ [bpcu]}$$

Remark: If we can't find $\underline{\mathbf{p}}$ that makes Y uniform, then $C < \log_2 |\mathcal{Y}| - H(\underline{\mathbf{r}})$ and we have to find a different technique to calculate C.



Example 4.35. Find the channel capacity of a BSC whose crossover probability is $p = 0.1$.

symmetric

$$\mathbf{Q} = \begin{array}{c} x \backslash y \\ 0 \\ 1 \end{array}\begin{array}{c} \begin{array}{cc} 0 & 1 \end{array} \\ \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix} \end{array} \quad \underline{r}$$

binary entropy func.

$$C = \log_2 |\mathcal{Y}| - H(\underline{r}) = \log_2 2 - H([0.9\ 0.1]) = 1 - H(\overset{p}{0.1}) \approx 0.531 \text{ [bpcu]}$$

Definition 4.36. A DMC is called **weakly symmetric** if (1) all the rows of its probability transition matrix $\mathbf{Q}$ are permutations of each other and (2) all the column sums are equal.

- It should be clear from the definition that a symmetric channel is automatically weakly symmetric.

Condition ① makes all $H(Y|X=x)$ the same. So, $H(Y|X)$ is easy to find and does not depend on $p(x)$

Condition ② guarantees that uniform X gives uniform Y.

Calculating channel capacity

c

1. Use (multi-variable) calculus
   - standard nonlinear optimization techniques
2. Use Blahut-Arimoto algorithm (MATLAB)
3. Check whether we can match the **Q** matrix with any known special cases.

Remark: Do not assume that the input probabilities will have to be uniform to obtain C.

- See BAC in Ex. 4.25.

Channel Capacity: Special Cases

- **Channel with Nonoverlapping Outputs (NO^2)**
  - There is only one non-zero element in each column of its **Q** matrix.
  - $\boxed{C = \log_2 |\mathcal{X}|}$ [4.30]
    is achieved by uniform input probabilities.
  - Ex. Noiseless Binary Channel: $C = 1$ [bpcu] [Ex. 4.27]
- **Weakly Symmetric Channel**
  - (1) all the rows of **Q** are permutations of each other and (2) all the column sums are equal. [Defn 4.36]
  - $\boxed{C = \log_2 |\mathcal{Y}| - H(\underline{\mathbf{r}})}$ where $\underline{\mathbf{r}}$ is any row from the **Q** matrix. [4.37]
    is achieved by uniform input probabilities.
  - Ex. Binary Symmetric Channel: $C = 1 - H(p)$ [bpcu]

ECS 452: In-Class Exercise # 11

Instructions

1. Separate into groups of no more than three students each. **The group cannot be the same as any of your former groups.** Only one submission is needed for each group.
2. **[ENRE] Explanation is not required for this exercise**
3. **Do not panic.**

Date: 28 / 2 / 2019			
Name	ID (last 3 digits)		

1. For each of the following DMC's probability transition matrices **Q**, (i) indicate whether the corresponding DMC is symmetric (Yes or No), (ii) indicate whether the corresponding DMC is weakly symmetric (Yes or No), (iii) evaluate the corresponding capacity value (your answer should be of the form X.XXXX).

To be symmetric DMC, the Q matrix must satisfy two conditions:
❶ rows are permutations of each other
❷ columns are permutations of each other

To be weakly symmetric DMC, the Q matrix must satisfy two conditions:
❶ rows are permutations of each other
❷ same column sum

The (informational) channel capacity is

$$C = \max_{\underline{\mathbf{p}}} I(\underline{\mathbf{p}}, \mathbf{Q})$$

However, in class, we have seen that there are multiple special cases which have easier formulas.

Q	Symmetric Channel?	Weakly Symmetric Channel?	C
$\begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix}$ $\begin{array}{cc} \downarrow\Sigma & \downarrow\Sigma \\ 1 & 1 \end{array}$	❶ ✓ ❷ ✓ Yes	❶ ✓ ❷ ✓ Yes	The DMC is a BSC with $p = 0.25$. $C = 1 - H(p) = 1 - H(0.25) \approx 1 - 0.8113 = 0.1887$.
$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5 & 0.5 \end{bmatrix}$ $\begin{array}{ccc} \Sigma\downarrow & \Sigma\downarrow & \Sigma\downarrow \\ 1 & 0.5 & 0.5 \end{array}$	❶ × ❷ × No	❶ × ❷ × No	The DMC satisfies the condition for NO². (Each column of **Q** has only one non-zero element.) $C = \log_2 \lvert\mathcal{X}\rvert = \log_2 2 = 1.000$.
$\begin{array}{c\|ccc} x \backslash y & y_1 & y_2 & y_3 \\ \hline x_1 & 0.25 & 0.5 & 0.25 \\ x_2 & 0.25 & 0.25 & 0.5 \\ x_3 & 0.5 & 0.25 & 0.25 \\ & \Sigma\downarrow & \Sigma\downarrow & \Sigma\downarrow \end{array}$	❶ ✓ ❷ ✓ Yes	❶ ✓ ❷ ✓ Yes	The DMC is weakly symmetric. $C = \log_2 \lvert\mathcal{Y}\rvert - H(\underline{\mathbf{r}}) = \log_2 3 - H\left(\left[\frac{1}{4}\ \frac{1}{2}\ \frac{1}{4}\right]\right) \approx 1.5850 - 1.5 = 0.0850$.
$\begin{array}{c\|cccc} x \backslash y & y_1 & y_2 & y_3 & y_4 \\ \hline x_1 & 0.75 & 0.25 & 0 & 0 \\ x_2 & 0 & 0.25 & 0 & 0.75 \\ x_3 & 0 & 0.25 & 0.75 & 0 \\ & \Sigma\downarrow & \Sigma\downarrow & \Sigma\downarrow & \Sigma\downarrow \\ & 0.75 & 0.75 & 0.75 & 0.75 \end{array}$	❶ ✓ ❷ × No	❶ ✓ ❷ ✓ Yes	The DMC is weakly symmetric. $C = \log_2 \lvert\mathcal{Y}\rvert - H(\underline{\mathbf{r}}) = \log_2 4 - H\left(\left[\frac{3}{4}\ \frac{1}{4}\right]\right) \approx 2 - 0.8113 = 1.1887$.

$$H\left(\left[\frac{1}{4}\ \frac{1}{2}\ \frac{1}{4}\right]\right) = -2 \times \frac{1}{4}\log_2\frac{1}{4} - \frac{1}{2}\log_2\frac{1}{2} = 1 + \frac{1}{2} = \frac{3}{2}.$$

$H(p) \equiv H([1-p \;\; p])$

$H(0.1) \equiv H([0.9 \;\; 0.1]) = H([0.1 \;\; 0.9]) = -0.1\log_2 0.1 - 0.9\log_2 0.9$

$-0.1\log_2 0.1 - 0.2\log_2 0.2 - 0.7\log_2 0.7 = H([0.1 \;\; 0.2 \;\; 0.7]) \neq H(0.1)$

4.37. For a weakly symmetric channel,

$$C = \log_2 |\mathcal{Y}| - H(\underline{\mathbf{r}}),$$

where $\underline{\mathbf{r}}$ is any row from the $\mathbf{Q}$ matrix. The capacity is achieved by a uniform pmf on the channel input.

- Important special case: For BSC, $C = 1 - H(p)$. crossover probability p

4.38. Properties of channel capacity

(a) $C \geq 0$

(b) $C \leq \min\{\log_2 |\mathcal{X}|, \log_2 |\mathcal{Y}|\}$

Example 4.39. Find the channel capacity of a DMC whose

$$\mathbf{Q} = \begin{bmatrix} 0.9 & 0.05 & 0.05 \\ 0.05 & 0.9 & 0.05 \\ 0.025 & 0.025 & 0.95 \end{bmatrix}$$

Suppose four choices are provided:
(a) 1.0944 (b) 1.5944 (c) 2.0944 (d) 2.5944

Example 4.40. Another case where capacity can be easily calculated: Find the channel capacity of a DMC of which all the rows of its $\mathbf{Q}$ matrix are the same.

$$Q = \begin{bmatrix} 1-\alpha & \alpha \\ 1-\alpha & \alpha \end{bmatrix} \quad \text{or} \quad Q = \begin{bmatrix} \boxed{\underline{r}} \\ \vdots \\ \boxed{\underline{r}} \end{bmatrix}$$

$H(Y|x) \equiv H(\underline{r}) \;\; \forall x \;\Rightarrow\; H(Y|X) = H(\underline{r})$

$\underline{q} = \underline{p}Q = [p_1 \;\; p_2 \;\; \cdots] \begin{bmatrix} \underline{r} \\ \underline{r} \\ \vdots \end{bmatrix}$

$= p_1\underline{r} + p_2\underline{r} + \cdots = \underline{r}\,\underbrace{\left(\sum_x p(x)\right)}_{1} = \underline{r}$

Therefore, $H(Y) = H(\underline{r})$ and

$I(X;Y) = H(Y) - H(Y|X) = H(\underline{r}) - H(\underline{r}) = 0$

regardless of the input distribution.

Another way to see this is that $Q(y|x)$ does not depend on x so $X \perp\!\!\!\perp Y$ and therefore $I(X;Y) = 0$ regardless of the input distribution.

$C = \max_{\underline{p}} \underbrace{I(X;Y)}_{0}$

$C = 0$

4.41. In this section, we worked with "toy" examples in which finding capacity is relatively easy. In general, there is no closed-form solution for computing capacity. When we have to deal with cases that do not fit in any special family of $\mathbf{Q}$ described in the examples above, the maximum may be found by standard nonlinear optimization techniques or the Blahut-Arimoto Algorithm discussed in 4.26.